Genetic mutations you want

To cure disease, researchers are starting to scour the genomes of the abnormally healthy.

Sarah C. P. Williams, *Science Writer*

In 2009, researchers at the Broad Institute in Boston, led by geneticist David Altschuler, started recruiting elderly, overweight individuals who, by all accounts, ought to have type 2 diabetes but didn't. The scientists weren't looking for genetic mutations that cause diabetes but rather hoping to find mutations that prevent it. Their search paid off; last year, the group reported in *Nature Genetics* that people who have particular mutations in a gene called *SLC30A8* (Solute carrier family 30, member 8) are 65% less likely to get diabetes, even when they have risk factors like obesity (1).

The gene has subtle effects on insulin, and, for a fortunate few, mutations that knock out its function seem to offset the forces that would, for the rest of us, likely lead to diabetes. Similarly protective mutations—that disable a gene but create a benefit rather than a problem—have been discovered somewhat accidentally in the past. One percent of Northern Europeans, for instance, are now known to carry a mutation in a gene called *CCR-5* that renders a cellular receptor defective and confers total immunity from HIV infection (2).

And there's evidence of more lucky mutations lurking in human genomes, in the form of people who seem to defy the odds—the long-lived smokers (3), or the individuals who remain unscathed in the midst of an infectious disease outbreak. Especially intriguing are those who carry gene mutations that are known to cause disease yet who show no signs of illness.

Now, cheaper sequencing is making it possible to hunt for these fairy godmother mutations and paving a more direct route toward turning discoveries into potential medications, or even targets for new gene editing techniques. It's a potentially fruitful strategy. Figuring out how to mimic the effects of a beneficial mutation is often simpler than determining how to reverse the effects of a detrimental one, says cardiologist and geneticist Sekar Kathiresan, also of the Broad Institute. "The most useful genetic findings are those that decrease a gene's function and protect against disease," he says. "These immediately tell you that if you can develop a drug that mimics the mutation, it should work in humans."

Finding these beneficial mutations, however, can be harder than finding disease-linked DNA changes. Recruiting people who rarely use the healthcare system is one hurdle. Another is that existing genetic databases are not usually designed to identify the absence of illness. But forging ahead despite these challenges is worthwhile, says Leslie Biesecker of the National Human Genome Research Institute (NHGRI). Scientists have long studied single nucleotide polymorphisms (SNPs) that are associated with disease, and investigating the opposite phenomenon will shed further light on the basic biology of how genes interact with one another, he says.

"We've been studying disease cohorts for a long time, and we've learned a lot from that. But if you really want to understand the full spectrum of the relationship between genes and disease, you have to study as many different kinds of people as you possibly can," says Biesecker. "You have to study diseased people, but you also have to study healthy people."

Beneficial mutations found in the "wellderly" or in disease survivors may point the way toward therapeutics. Image courtesy of Dave Cutler.

The Unusually Well

Because so many chronic illnesses don't manifest until later in life, the unusually healthy elderly are one good

www.pnas.org/cgi/doi/10.1073/pnas.1601663113

place to start a search for protective mutations. There have been a few hypotheses about why some people live such long, healthy lives, says Nir Barzilai of Albert Einstein College of Medicine. "One was that these guys have the perfect genome; they just don't have any of the mutations that are associated with disease," he says. "Another was that they're all lean, nonsmoking vegetarians."

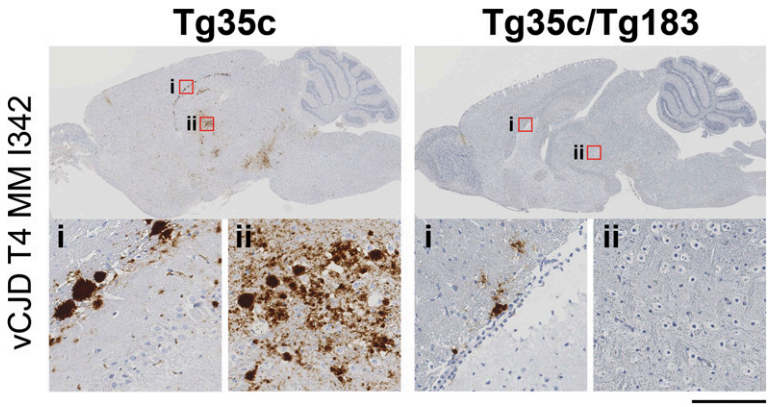
Recent studies quash both these theories. Last year, Barzilai's group analyzed 44 full genome sequences from centenarians. In total, the group had 250 mutations linked to Parkinson's, Alzheimer's, cardiovascular disease, and other chronic conditions, the scientists reported in *Molecular Genetics and Genomic Medicine* (4). Moreover, some of the hundred-year-olds were obese, others had been lifelong smokers, and many had never regularly exercised. However, they'd all lived a century, and none had developed signs of chronic disease. "That leaves us with the fact that they must have some genomic reasons—other than the lack of disease genes—for their longevity," says Barzilai.

Shortly after Barzilai's study was published, NHGRI researchers led by Biesecker analyzed the protein-coding genes, or exomes, of 951 healthy adults and found that 1 in 10 had mutations linked to Parkinson's, heart defects, and blood disorders, among other things. These were gene variants that don't just increase disease risk but are thought to always cause disease. But half of those people were not ill (5).

Despite such tantalizing clues, searches for the protective mutations that could be offsetting the effects of disease-linked genes and lifestyle factors have been hit and miss so far. In 2007, Eric Topol of the Scripps Institute and his colleagues, eager to look at a concentrated collection of healthy genomes, began recruiting people over the age of 80 who didn't have chronic diseases and weren't on medications, as part of the Scripps Wellderly Project. Over the next 7 years, they developed a cohort of 1,400 so-called wellderly. In 2014, they published the full genomes of 454 participants in an open-access database for researchers anywhere in the world to use. So far, no protective mutations have been turned up. But "the hunt is on," Topol says.

Over a similar period, Barzilai, keen to focus on a relatively homogeneous population to facilitate the discovery of genetic variants, studied Ashkenazi Jews over the age of 95. Barzilai's LonGenity Project has collected genetic and health information from over 500 of these extreme elderly as well as 700 of their offspring. Even before they'd completed full genomes of their centenarians, Barzilai and his colleagues had turned up two promising gene variants. A deletion in the adiponectin gene *ADIPOQ*, they found, appeared to protect against inflammation of arteries (6). And a mutation in the cholesteryl ester transfer-protein gene *CETP* was seen more often in the older cohort, and was linked to protection against both high cholesterol levels and cognitive disorders (7).

Other researchers, rather than recruiting their own healthy cohorts, have sifted through existing databases, such as that of the National Heart, Lung, and Blood Institute's ongoing Grand Opportunity (GO) Exome Sequencing Project. It includes samples from some 200,000 participants in large population health studies like the multigenerational Framingham Heart Study. Katherisan's group mined a hundred thousand exomes from the GO collection to discover gene variants associated with low cholesterol levels (8). And scientists at the commercial genome sequencing company 23andMe announced that they had used their growing database to find that people with mutations in the gene *SGK1* are less likely than most to develop Parkinson's, even if they have risk factors for the disease.



One example of the effects of beneficial mutations. (*Left*) Postmortem mouse brains show abnormal prion deposition, specifically in the hippocampus (*i*) and thalamus (*ii*). (*Right*) However, mice with protective genes show only weak levels of prion deposition, including in the thalamus (*ii*). Adapted from ref. 13 with permission from Macmillan Publishers Ltd: *Nature*, copyright 2015.

Pathways to Therapeutics

The wellderly aren't the only ones harboring potential genetic gems; examining infectious disease survivors offers another promising avenue. "Wherever there's been a profound infectious disease infecting a community, looking at the survivors enables you to look for resistance genes which may cast enormous light on the etiology of the disease and potentially lead to new treatment," says neurologist John Collinge of University College London.

Researchers, for example, are investigating drugs to fight the Ebola virus that target a protein known as Niemann Pick type C (NPC). The gene that encodes it, when mutated, causes a rare version of Niemann Pick disease that is usually fatal in childhood to people with two copies. But in animal studies, individuals with only one mutated copy of the gene resist Ebola infection because the virus needs the working version of the protein to infect host cells (9).

In other recent research, investigators have looked for gene mutations that protect against infection, or severe illness from influenza and other pathogens (10). In October, researchers with the MalariaGEN international consortium identified a gene variant that affects a blood cell surface receptor and protects against severe cases of malaria (11).

Collinge and his colleagues have been studying survivors of a more exotic epidemic: kuru, a deadly neurological illness similar to Creutzfeld Jakob Disease (CJD). Like CJD, kuru is transmitted by

misshapen proteins called prions, and, in the 1950s and 1960s, it spread rapidly among members of a cannibalistic tribe in Papa New Guinea. When someone died of kuru, ritualistic consumption of their body meant that those participating in the ceremony would contract the disease too. In some villages, almost all of the women of childbearing age perished.

But decades later, there were also survivors—people who had partaken of the feasts and never gotten sick. In the early 1990s, Collinge began sequencing their genomes. Over the past two decades, he's revealed mutations in their prion protein gene, *PRNP*, that protect them from kuru (12).

"In those families with the polymorphism, there's hardly any kuru despite very high levels of exposure," says Collinge. This year, Collinge and his colleagues reported in *Nature* that mice with one of the mutations were protected from 18 different kinds of prion disease (13). "This particular finding is incredibly powerful," says Collinge. "We went from 100 percent of the mice dying to 0 percent." Now, the researchers are working on determining the structure of the protective prion proteins, which could shed light on how to mimic the mutation in the rest of the human population, possibly leading to treatments for not just kuru but a variety of prion diseases.

Risk and protection are really just flip sides of the same coin.

—Sekar Kathiresan

Ideally, the discovery of a protective mutation can inform the development of a drug that mimics its molecular effects in the body. Inhibitors of CETP, studied by Barzilai, have been explored as cholesterol drugs, although none has reached the market. And 23andMe's discovery that some *SGK1* mutations protect against Parkinson's has been followed up with basic research showing that blocking SGK1, a protein known to mediate the way cells respond to stress, can turn off pathways involved in neurodegeneration (14). The advent of Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)/Cas9 gene editing technology offers still more possibilities for developing therapies based on beneficial gene mutations. The controversial technique could one day provide a way to alter the genes of adults for the better. "It is imaginable that in addition to fixing disease-causing mutations, CRISPR/Cas9 will be used to make changes to genes that lower the risk for disease," says biologist Jonathan Weissman of the University of California, San Francisco, whose research includes CRISPR applications.

The ability to find natural protective mutations might even speed the progress of drug testing—a process that's typically slow and expensive—by helping to validate drug targets. Such clues could, in principle, help drug companies decide on the drugs most likely to be effective. In an experiment to predict whether a compound was likely to have the desired effect, Kathiresan and colleagues set out to see if a drug called ezetimibe—developed to lower cholesterol—would also prevent heart attacks. Because ezetimibe blocks the NPC1L1 protein, the team looked for people with mutations in the *NPC1L1* gene to study their heart attack rates, "If we could find these people, it would be as if they'd be given the drug for their whole life," says Kathiresan.

Returning to a subset of the NHLBI's hundred thousand exomes, Kathiresan's team found a handful—roughly one in 650 people—who had any of 15 *NPC1L1* mutations and, indeed, those individuals had a 53% lower heart attack risk compared with people without the mutations. A few months later, the results of a clinical trial came back; advanced heart disease patients taking ezetimibe showed a small decrease in heart attacks and strokes (15). It was proof of concept that beneficial mutations could help predict the effect of a drug.

A Struggle Against Statistics

Kathiresan's experiment depended on the huge NHLBI exome database because beneficial mutations are both hard to find and hard to prove. For his plan to predict drugs' performance in trials, as with any efforts to hunt down protective mutations, researchers need very large pools of people and loads of data on their health.

If a few people with a rare disease also all share a rare genetic mutation, there's a good bet that the mutation is related to their disease. But if a handful of healthy people have the same genetic mutation, it's more likely to be coincidence, and more difficult—from a statistical standpoint—to demonstrate causation.

"Risk and protection are really just flip sides of the same coin," says Kathiresan. "If you have a mutation that increases risk in 5 percent of people, you could really say that 95 percent of people have a protective version of the gene." When he and his team looked for mutations linked to low blood triglycerides, they decided their quarry had to both knock out or impede a protein's function and lower risk below the norm. Amid 100,000 exomes, they managed to find four variants in *APOC3*, each of which occurs in only around 1 in 1,000 people.

Biesecker's ClinSeq study, with under 1,000 participants, isn't even designed to seek out protective mutations, only to document examples of people with disease-causing gene variants but no disease. That's because getting enough people to search for disease-preventing genes is such a challenge, Biesecker says. Complicating matters, vast networks of related genes might contribute to a given disease or set of symptoms.

"We've long known that you can have gene–gene interactions and that one gene variant can compensate for another. But these things are statistically and mathematically challenging to study because the combinatorial possibilities here are enormous," he says. "It's a numbers and power issue. We'd need millions of people in a cohort to be able to statistically tease those things out."

Efforts are underway to construct enormous databases that can be mined for protective mutations. Perhaps the most ambitious is the Resilience Project, led by researchers at the Icahn School of Medicine at Mount Sinai in New York and Sage Bionetworks in Seattle. They are attempting to solicit 1 million volunteers to donate DNA samples (16). The project's focus is finding people in this huge random sample who harbor gene mutations known to cause rare and severe disorders when a single gene copy is present, such as Costello syndrome and Cardiofaciocutaneous Syndrome, yet who may not even know they have a disease.

The potential value of such a database, especially the prospect of including detailed health histories to detect the presence or absence of illness, is illustrated by the lucky break that led Altschuler's group at the Broad group to zero in on one variant of the *SLC30A8* gene. The team had a data suggesting that *SLC30A8* might be protective, but they couldn't quite come up with the statistical power they needed to prove the gene's effect. In Iceland, however, neurologist Kari Steffanson and his company deCODE genetics has spent two decades compiling genetic and health data on half a million people, including more than a third of Icelanders. The trove includes 10,000 whole-genome sequences, 2,600 of which were described in a *Nature Genetics* paper last year (17). When one of the Broad scientists mentioned their suspicions about *SLC30A8* to Steffanson during a phone call, the deCODE CEO did a quick search through his database for people who had the mutation—and their health backgrounds. "They had a hint of an association but could never prove it," Steffanson says. "Within 10 minutes, though, I could demonstrate that we had variants."

Still, Steffanson, too, thinks more genomes and more phenotype information are needed. "There is the role of chance, there is the role of the environment, and there is the role of the rest of the genetic background," Steffanson points out. "So this is a complex interplay."

"More people are doing this kind of study now," says Barzilai. "But, unfortunately, not enough."

Studying the extremely elderly or extremely healthy, he says, has the potential to help researchers make connections between genes and their function, between diseases and their molecular causes, and between therapeutics and their effectiveness. "If studies like ours are successful," says Barzilai, "we can profoundly change both aging and disease."

1 Flannick J, et al.; Go-T2D Consortium; T2D-GENES Consortium (2014) Loss-of-function mutations in *SLC30A8* protect against type 2 diabetes. *Nat Genet* 46(4):357–363.
2 Liu R, et al. (1996) Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. *Cell* 86(3):367–377.
3 Levine ME, Crimmins EM (2015) A genetic network associated with stress resistance, longevity, and cancer in humans. *J Gerontol Ser A*, in press.
4 Freudenberg-Hua Y, et al. (2014) Disease variants in genomes of 44 centenarians. *Mol Genet Genomic Med* 2(5):438–450.
5 Johnston JJ, et al. (2015) Individualized iterative phenotyping for genome-wide analysis of loss-of-function mutations. *Am J Hum Genet* 96(6):913–925.
6 Barzilai N, et al. (2010) Genetic studies reveal the role of the endocrine and metabolic systems in aging. *J Clin Endocrinol Metab* 95(10):4493–4500.
7 Schechter CB, Barzilai N, Crandall JP, Atzmon G (2010) Cholesteryl ester transfer protein (CETP) genotype and reduced CETP levels associated with decreased prevalence of hypertension. *Mayo Clin Proc* 85(6):522–526.
8 Stitziel NO, et al.; Myocardial Infarction Genetics Consortium Investigators (2014) Inactivating mutations in *NPC1L1* and protection from coronary heart disease. *N Engl J Med* 371(22):2072–2082.
9 Côté M, et al. (2011) Small molecule inhibitors reveal Niemann-Pick C1 is essential for Ebola virus infection. *Nature* 477(7364):344–348.
10 Hsu T-H, Spindler KR (2012) Identifying host factors that regulate viral infection. *PLoS Pathogens* 8(7):e1002772.
11 Band G, Rockett KA, Spencer CC, Kwiatkowski DP; Malaria Genomic Epidemiology Network (2015) A novel locus of resistance to severe malaria in a region of ancient balancing selection. *Nature* 526(7572):253–257.
12 Mead S, et al. (2009) A novel protective prion protein variant that colocalizes with kuru exposure. *N Engl J Med* 361(21):2056–2065.
13 Asante EA, et al. (2015) A naturally occurring variant of the human prion protein completely prevents prion disease. *Nature* 522(7557): 478–481.
14 Iqbal S, et al. (2015) Serum- glucocorticoid-inducible kinase 1 confers protection in cell-based and in in vivo neurotoxin models via the c-Jun N-terminal kinase signaling pathway. *Mol Cell Biol* 35(11):1992–2006.
15 Cannon CP, et al.; IMPROVE-IT Investigators (2015) Ezetimibe added to statin therapy after acute coronary syndromes. *N Engl J Med* 372(25):2387–2397.
16 Friend SH, Schadt EE (2014) Translational genomics. Clues from the resilient. *Science* 344(6187):970–972.
17 Gudbjartsson DF, et al. (2015) Large-scale whole-genome sequencing of the Icelandic population. *Nat Genet* 47(5):435–444.